



US Army Corps
of Engineers®

Bayesian Inference of Nonstationary Precipitation Intensity-Duration- Frequency Curves for Infrastructure Design

*by Brian E. Skahill, Amir AghaKouchak, Linyin Cheng,
Aaron Byrd, and Joseph Kanney*

PURPOSE: The purpose of this document is to demonstrate the application of Bayesian Markov Chain Monte Carlo (MCMC) simulation as a formal probabilistic-based means by which to develop local precipitation Intensity-Duration-Frequency (IDF) curves using historical rainfall time series data collected for a given surface network station, including the treatment of a nonstationary climate condition. This objective will be accomplished by independently revisiting parts of an example originally profiled by Cheng and AghaKouchak (2014). This Technical Note will conclude with a brief discussion of some potential opportunities for future U.S. Army Corps of Engineers (USACE) research and development directed at extreme rainfall frequency analysis.

INTRODUCTION: The rainfall IDF curve is a mathematical relationship between rainfall intensity, duration, and return period (Koutsoyiannis et al. 1998). IDF curves describe rainfall intensity as a function of duration for a given return period and are important for infrastructure design (Overeem et al. 2008). The construction of IDF curves involves fitting a theoretical distribution to the historical extreme rainfall amounts for a number of fixed durations (Overeem et al. 2008). Current IDF curves are developed based on the concept of temporal stationarity, which assumes that the occurrence probability of extreme precipitation events is not expected to change significantly over time.

The Fifth Assessment Report (AR) of the Intergovernmental Panel on Climate Change (IPCC 2013) reported global surface temperatures to increase 0.3 to 4.8 degrees Celsius ($^{\circ}\text{C}$) by the year 2100 relative to the reference period 1986–2005 (Srivastav et al. 2014). The Fourth AR of the IPCC (IPCC 2007) reported that the global surface temperature had increased approximately 0.75 $^{\circ}\text{C}$ over the last 100 years (the increase could not be explained by natural variability alone) and moreover, that increased greenhouse gas emissions due to human activities are the main reason for current global warming (Yilmaz et al. 2014). With every 1 $^{\circ}\text{C}$ warming of the global surface temperature, the atmosphere's water holding capacity is increased by approximately 7%. Hence, global warming directly affects a changing precipitation climatology (Trenberth 2011), including its extremes (Kunkel et al. 2013), and this change is termed *nonstationarity*. Kunkel et al. (2013) concluded, based on climate model simulations and conceptual considerations of other related meteorological systems, that probable maximum precipitation values will increase in the future due to increased atmospheric moisture content and higher levels of moisture transport into storms.

Cheng and AghaKouchak (2014) outlined a general methodology for developing IDF curves under nonstationary conditions. Their framework involved application of the Generalized Extreme Value (GEV) distribution, annual rainfall maxima data associated with a recent NOAA

(National Oceanic and Atmospheric Administration) Atlas 14 update (see Perica et al. 2013 and references cited therein), and a Bayesian MCMC sampler for simultaneous optimization and inference (ter Braak 2006). Nonstationarity was treated by defining the location parameter of the GEV distribution to vary linearly with time. To illustrate the potential negative impacts of ignoring nonstationarity at a site that has been assessed otherwise, by way of identification of a statistically significant increasing trend in precipitation extremes, Cheng and AghaKouchak (2014) applied their framework to develop IDF curves at five distinct rainfall station locations with and without treatment of a nonstationary climate condition. In this USACE Engineer Research and Development Center (ERDC) Coastal and Hydraulics Laboratory (CHL) Technical Note (TN), the previously mentioned Bayesian supervised IDF curve development analysis for the White Sands National Monument rainfall station is independently revisited, albeit by applying a different MCMC sampler (ter Braak and Vrugt 2008), not only to demonstrate related internal USACE-ERDC-CHL capacity but also to further focus on a comparison of these Bayesian-inferred IDF curves under stationary and nonstationary conditions.

BACKGROUND: Probability concepts and related relevant terms in hydrology are summarized in Stedinger et al. (1992). For a random variable X (e.g., herein, representing the annual rainfall maxima), its cumulative distribution function (cdf), denoted by $F_X(x)$, is the probability that the random variable X is less than or equal to x :

$$F_X(x) = P(X \leq x) \quad (1)$$

The probability density function (pdf) defines the relative likelihood that X takes on different values. It is denoted by $f_X(x)$, and it is the first derivate of the cumulative distribution function. In hydrology, the p th quantile, denoted by x_p , is the value with cumulative probability p :

$$F_X(x_p) = p \quad (2)$$

In addition, the return period associated with the p th quantile x_p , denoted and defined by $T = 1/(1-p)$, represents the average frequency of occurrence for an event of magnitude x_p .

In this study, the GEV distribution is assumed:

$$f_X(x) = \frac{1}{\sigma} \left[1 - \frac{\xi(x-\mu)}{\sigma} \right]^{1/\xi-1} \exp \left\{ - \left[1 - \frac{\xi(x-\mu)}{\sigma} \right]^{1/\xi} \right\} \quad (3)$$

$$F_X(x) = \begin{cases} \exp \left\{ - \left[1 - \frac{\xi(x-\mu)}{\sigma} \right]^{1/\xi} \right\} & \xi \neq 0 \\ \exp \left\{ - \exp \left[- \frac{(x-\mu)}{\sigma} \right] \right\} & \xi = 0 \end{cases} \quad (4)$$

$$x_p = \begin{cases} \mu + \frac{\sigma}{\xi} \left[1 - (-\ln(p))^\xi \right] & \xi \neq 0 \\ \mu - \sigma \ln(-\ln(p)) & \xi = 0 \end{cases} \quad (5)$$

where $\mu + \sigma / \xi \leq x < \infty$ for $\xi < 0$, $-\infty < x < \infty$ for $\xi = 0$, $-\infty < x \leq \mu + \sigma / \xi$ for $\xi > 0$, and μ, σ , and ξ are the location, scale, and shape parameters of the distribution. For the GEV distribution, the p th quantile defined in Equation 5 is obtained by inverting Equations 2 and 4.

The MCMC simulation is a formal Bayesian approach for estimating the posterior probability distribution of the specified adjustable model parameters, in this case, the GEV distribution parameters. It treats the specified adjustable model parameters as random variables and relies upon Bayes' Theorem to compute their joint posterior probability distribution. Bayes' Theorem effectively communicates that the posterior distribution is proportional to the product of the prior distribution and the likelihood function (i.e., the conditional distribution), which encapsulates the conditioning process with the observed dataset. The prior distribution is prescribed based on the modeler's best judgment, expert opinion, or literature estimates, among possible others. In this case, the observed dataset is a systematic record of annual rainfall maxima at the White Sands National Monument rainfall station. The idea behind MCMC simulation is that while one wants to compute a probability density, $p(\mathbf{p}|D)$, where \mathbf{p} and D represent the vector of adjustable model parameters and the data/information imparted to the analysis, respectively, there is the understanding that such an endeavor may be impracticable. Additionally, simply being able to generate a large random sample from the probability density would be equally sufficient as knowing its exact form. Hence, the problem then becomes one of effectively and efficiently generating a large number of random draws from $p(\mathbf{p}|D)$. It was discovered that an efficient means to this end is to construct a Markov chain, a stochastic process of values that unfold in time, with the following properties: (1) the state space (set of possible values) for the Markov chain is the same as that for \mathbf{p} ; (2) the Markov chain is easy to simulate from; and (3) the Markov chain's equilibrium distribution is the desired probability density $p(\mathbf{p}|D)$. The Gelman and Rubin (1992) quantitative measure is commonly employed to assist with diagnosis of chain convergence. A Markov chain with the aforementioned properties can be constructed by choosing a symmetric proposal distribution and employing the Metropolis acceptance probability (Metropolis et al. 1953) to accept or reject candidate points. By constructing such a Markov chain, one can then simply run it to equilibrium (and this period is often referred to as the sampler *burn-in* period) and subsequently sample from its stationary distribution. Within the context of its application to simultaneously optimize and infer the GEV distribution parameters using a record of annual rainfall maxima for a given duration, the post burn-in random draws from p can be used to construct predictive distributions and credible intervals for particular quantiles by using Equation 5.

A key element for MCMC samplers that employ the Metropolis or Metropolis-Hastings acceptance probability rule (Metropolis et al. 1953; Hastings 1970) is the proposal distribution, which generates the candidate jumps for consideration as part of the Markov Chain directed random walk of the posterior. For a given problem, there are many possible acceptable proposal distributions. However, its specific choice can dramatically impact the overall efficiency of the sampler, to the target equilibrium distribution. Proposal distributions that generate either small or large jumps yield low acceptance rates and slow convergence. The primary goal is to choose a proposal

distribution that is easy to sample from, generates unbiased moves, and results in optimal mixing of the chains. The interested reader is directed to Gelman et al. (2004) for more information regarding technical details related to Bayesian MCMC.

As with Cheng and AghaKouchak (2014), nonstationarity is treated by defining the GEV location parameter to vary linearly in time, t :

$$\mu = \mu(t) = \mu_1 t + \mu_0 \quad (6)$$

This specific temporal treatment of the GEV location parameter is but one of many possible time variable approaches one could apply for IDF curve development using Bayesian MCMC. The linear in-time treatment of the GEV location parameter is employed primarily for the purposes of demonstration. In so doing, an additional random variable is introduced such that there are now four random variables, viz., μ_1, μ_0, σ , and ξ to simultaneously optimize and infer using MCMC. The quantiles are computed from the post burn-in random draws as described in Cheng and AghaKouchak (2014). In particular, for each given post burn-in random draw, the 95th percentile value for the location parameter, obtained by applying equation 6 (i.e., 95th percentile of $\mu(t=1), \dots, \mu(t=100)$), is used to compute its related quantile value for a specified value of p . The entire set of x_p computed from the post burn-in draws characterize its posterior predictive distribution.

EXAMPLE: Annual rainfall maxima series associated with the NOAA Atlas 14 update for the White Sands National Monument rainfall station located in the state of New Mexico (latitude: 32.7817; longitude: 106.1747; elevation: 1217.7 m) for the 52-year period 1949–2000 were collected using the National Oceanic Atmospheric Administration National Weather Service Hydrometeorological Design Studies Center Precipitation Frequency Data Server for eight specific durations (1 hr, 2 hr, 3 hr, 6 hr, 12 hr, 24 hr, 48 hr, and 96 hr) to support Bayesian supervised IDF curve development analysis under stationary and nonstationary conditions. In particular, 16 distinct MCMC simulations were performed to develop IDF curves under stationary and nonstationary conditions using the annual rainfall maxima for the noted eight durations for the White Sands National Monument rainfall station. As previously mentioned, an adaptive population-based MCMC sampler (ter Braak and Vrugt 2008) was employed to infer the joint posterior for the GEV distribution parameters. All 16 MCMC simulations specified a population size to evolve equal in value to five times the dimensionality of the estimation problem. Latin hypercube sampling was used to initialize the population. The size of the initial history of past simulated states to draw upon to generate jump proposals was specified to be equal in value to one hundred times the dimensionality of the estimation problem. For each simulation, an uninformed uniform prior distribution was employed as well as a likelihood function of the form

$$l(D|p) = \prod_{i=1}^s f_x(y_i|p) \quad (7)$$

where D is the sample set of recorded annual rainfall maxima for a given duration, y_i , s is the record size, and f is the GEV distribution. For each MCMC simulation, the Gelman and Rubin

(1992) quantitative convergence diagnostic was used together with visual inspection of trace plots of the chains and an efficiency plot of the evolving population mean root mean squared error to in aggregate assess the completion of sampler burn-in. In each case, subsequent to the weight of evidence-based assessment that sampling is occurring with stable frequency from the target distribution implied by the modeling analysis, a thinned history, viz., every tenth evolution of the approximately 1 million specified total post burn-in monitoring period model runs was saved and used to support IDF curve development.

Results from the 16 MCMC simulations are summarized in Tables 1–2 and Figures 1–8. Tables 1 and 2 list the posterior mode (PM) (i.e., the GEV with \mathbf{p} which maximizes $p(\mathbf{p}|D)$) estimates computed for each of the eight previously mentioned duration-based simulations under stationary and nonstationary conditions, respectively, and in each case, the related quantile estimates calculated for five distinct return periods, viz., 2, 10, 25, 50, and 100 years. It is underscored to the reader that the nonstationary results, as mentioned above, are computed and processed for a discrete point in time. Tables 1 and 2 also list by duration the computed 2.5, 50, and 97.5 percentiles of the posterior predictive distribution for each of the five quantiles. These noted percentile values are the basis for Figures 1–8, which are plots of the IDF curves computed under stationary and nonstationary conditions for the White Sands National Monument station, by duration. On each IDF curve shown in Figures 1–8, the 95% credible interval, based on the respective 2.5 and 97.5 percentile values listed in Tables 1 and 2, is shown for each quantile. Each Figure also includes plots of the predictive distributions derived from the post burn-in random draws for the 2-year and 100-year return period quantiles. Table 3 summarizes the computed percent increase obtained when the nonstationary, PM-based quantile estimates presented in Table 2 are compared with their counterparts listed in Table 1, which were obtained assuming a stationary climate. Figure 9 is a plot of the data presented in Table 3.

Table 1. Tabular summary by duration of the PM estimate for the GEV distribution parameters, its related computed quantiles, and also the 2.5, 50, and 97.5 percentile values from each respective predictive posterior distribution for five return periods computed under stationary conditions.									
Stationary Treatment									
Duration (hours)	PM			Quantiles (mm/h) for Return Periods, T in years; PM estimate, and Percentiles (2.5, 50, 97.5) of Predictive Distribution					
	μ	σ	ξ	2	10	25	50	100	
1	13.235	5.946	-0.025	15.425 (13.546, 15.564, 17.828)	27.008 (24.061, 27.853, 34.630)	33.051 (28.803, 34.219, 47.047)	37.630 (31.981, 39.061, 58.616)	42.256 (34.833, 43.964, 72.428)	
2	15.670	6.453	-0.068	9.033 (7.979, 9.108, 10.384)	15.678 (13.930, 16.204, 20.468)	19.357 (16.708, 20.106, 28.722)	22.244 (18.607, 23.183, 36.865)	25.248 (20.335, 26.412, 47.094)	
3	17.004	6.753	-0.082	6.506 (5.749, 6.556, 7.473)	11.231 (9.988, 11.621, 15.039)	13.900 (11.952, 14.479, 21.982)	16.018 (13.246, 16.776, 29.234)	18.245 (14.409, 19.216, 38.929)	
6	20.200	7.874	0.081	3.841 (3.407, 3.855, 4.338)	6.065 (5.549, 6.234, 7.490)	7.062 (6.391, 7.293, 9.733)	7.754 (6.900, 8.037, 11.755)	8.403 (7.315, 8.744, 14.096)	
12	22.163	8.534	0.128	2.102 (1.863, 2.103, 2.360)	3.238 (2.974, 3.311, 3.892)	3.714 (3.389, 3.811, 4.884)	4.032 (3.628, 4.146, 5.736)	4.321 (3.818, 4.455, 6.690)	
24	25.446	9.454	0.183	1.200 (1.074, 1.203, 1.341)	1.787 (1.659, 1.826, 2.098)	2.014 (1.866, 2.063, 2.543)	2.159 (1.981, 2.215, 2.901)	2.285 (2.071, 2.350, 3.285)	
48	28.541	11.790	0.055	0.684 (0.604, 0.687, 0.776)	1.114 (1.011, 1.145, 1.375)	1.315 (1.180, 1.358, 1.773)	1.457 (1.287, 1.509, 2.124)	1.592 (1.381, 1.654, 2.525)	
96	36.215	14.399	0.104	0.431 (0.383, 0.432, 0.485)	0.678 (0.622, 0.695, 0.815)	0.786 (0.716, 0.810, 1.017)	0.859 (0.775, 0.888, 1.186)	0.926 (0.825, 0.961, 1.370)	

Table 2. Tabular summary by duration of the PM estimate for the time varying GEV distribution parameters, its related computed quantiles, and also the 2.5, 50, and 97.5 percentile values from each respective predictive posterior distribution for five return periods computed under nonstationary conditions. The reported GEV location parameter is computed using the PM derived estimates for μ_1 and μ_0 , and Equation 6 for the 95th percentile.

Nonstationary Treatment									
Duration (hours)	PM			Quantiles (mm/h) for Return Periods, T in years; PM estimate, and Percentiles (2.5, 50, 97.5) of Predictive Distribution					
	μ	σ	ξ	2	10	25	50	100	
1	22.244	5.623	-0.009	24.308 (17.487, 24.917, 34.818)	35.028 (28.700, 36.610, 46.513)	40.494 (33.841, 42.546, 54.510)	44.580 (37.298, 46.821, 63.478)	48.661 (40.335, 50.909, 75.449)	
2	27.095	6.022	-0.039	14.659 (10.869, 14.876, 20.181)	20.629 (17.195, 21.434, 26.889)	23.803 (20.208, 24.946, 31.946)	26.235 (22.238, 27.543, 37.941)	28.716 (24.043, 30.082, 45.955)	
3	31.282	6.273	-0.030	11.198 (8.224, 11.285, 15.352)	15.295 (12.670, 15.779, 19.735)	17.447 (14.790, 18.165, 23.222)	19.083 (16.176, 19.882, 27.786)	20.741 (17.381, 21.528, 34.271)	
6	41.110	6.598	0.090	7.248 (5.671, 7.239, 8.966)	9.092 (7.708, 9.275, 11.072)	9.909 (8.560, 10.213, 12.376)	10.471 (9.093, 10.853, 13.752)	10.995 (9.539, 11.436, 15.524)	
12	45.233	7.133	0.149	3.981 (3.128, 3.994, 4.924)	4.906 (4.161, 5.002, 5.952)	5.282 (4.575, 5.426, 6.448)	5.529 (4.823, 5.704, 6.883)	5.749 (5.027, 5.950, 7.416)	
24	49.271	7.434	0.108	2.164 (1.758, 2.180, 2.641)	2.671 (2.301, 2.725, 3.209)	2.890 (2.508, 2.959, 3.528)	3.038 (2.630, 3.115, 3.827)	3.175 (2.725, 3.254, 4.197)	
48	53.067	9.410	-0.070	1.178 (0.952, 1.191, 1.461)	1.583 (1.356, 1.630, 1.971)	1.808 (1.544, 1.869, 2.413)	1.985 (1.669, 2.050, 2.903)	2.169 (1.779, 2.235, 3.540)	
96	59.863	12.945	0.052	0.673 (0.495, 0.672, 0.859)	0.910 (0.742, 0.932, 1.147)	1.021 (0.846, 1.054, 1.351)	1.100 (0.912, 1.139, 1.556)	1.176 (0.968, 1.218, 1.807)	

Table 3. Computed percent increase obtained when the nonstationary PM-based quantile estimates are compared with their counterparts that were computed assuming a stationary climate.

Duration (hours)	Percent Increase in Quantiles (mm/h) for Return Periods, T in years				
	2	10	25	50	100
1	57.6	29.7	22.5	18.5	15.2
2	62.3	31.6	23.0	17.9	13.7
3	72.1	36.2	25.5	19.1	13.7
6	88.7	49.9	40.3	35.0	30.9
12	89.5	51.5	42.2	37.1	33.1
24	80.4	49.5	43.5	40.7	38.9
48	72.3	42.1	37.5	36.2	36.2
96	56.0	34.2	30.0	28.1	27.0

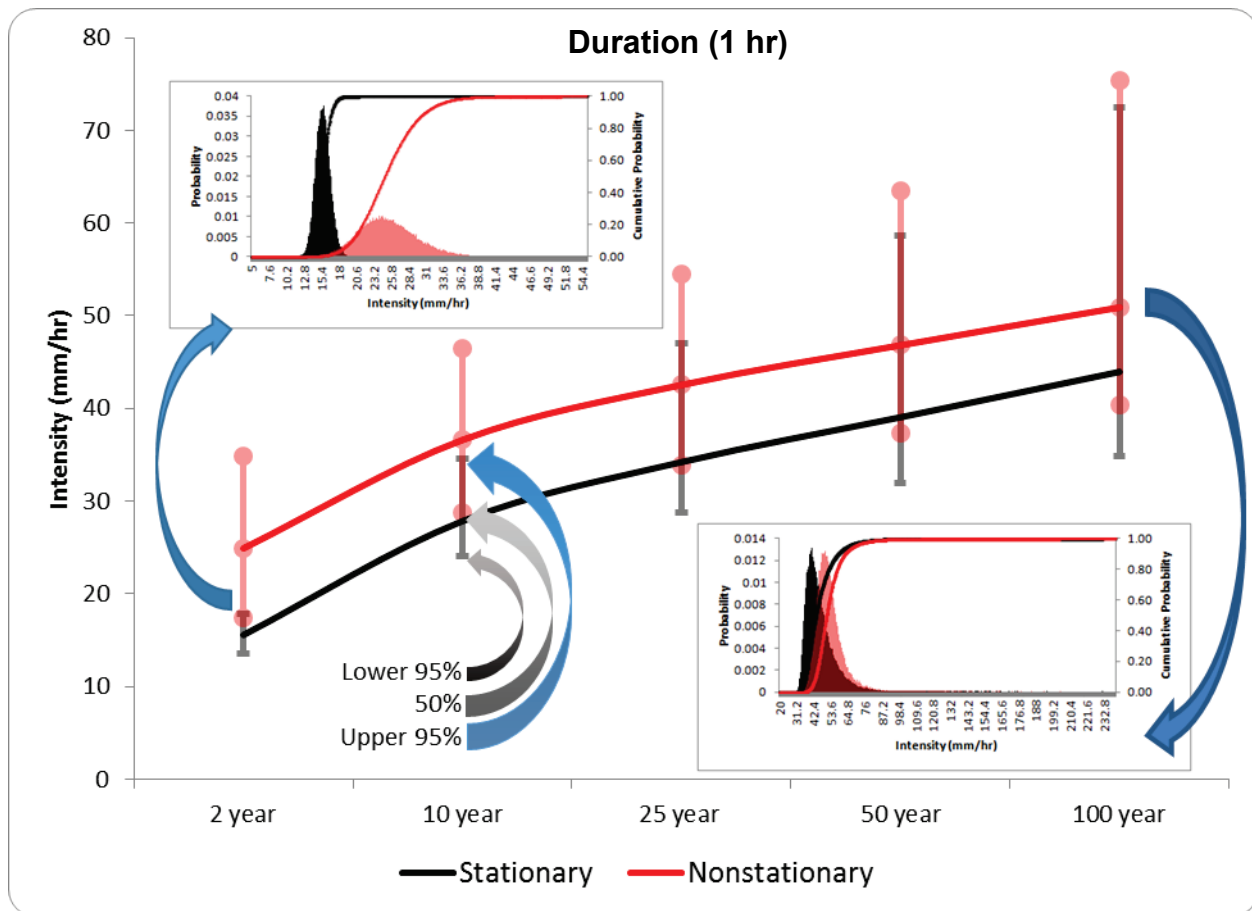


Figure 1. Bayesian MCMC simulation-derived, 1 hr IDF curves for the White Sands National Monument rainfall station computed under stationary and nonstationary conditions. The 2.5, 50, and 97.5 percentile values from each respective predictive posterior distribution are shown at each quantile level. These three values are clearly identified for the stationary 10-year return period results. Their relative locations equally apply for the remaining return periods for both the stationary and nonstationary analyses not only in this figure but also Figures 2–8. Plots of the posterior predictive distributions for the 2-year and 100-year return period level quantiles are also shown.

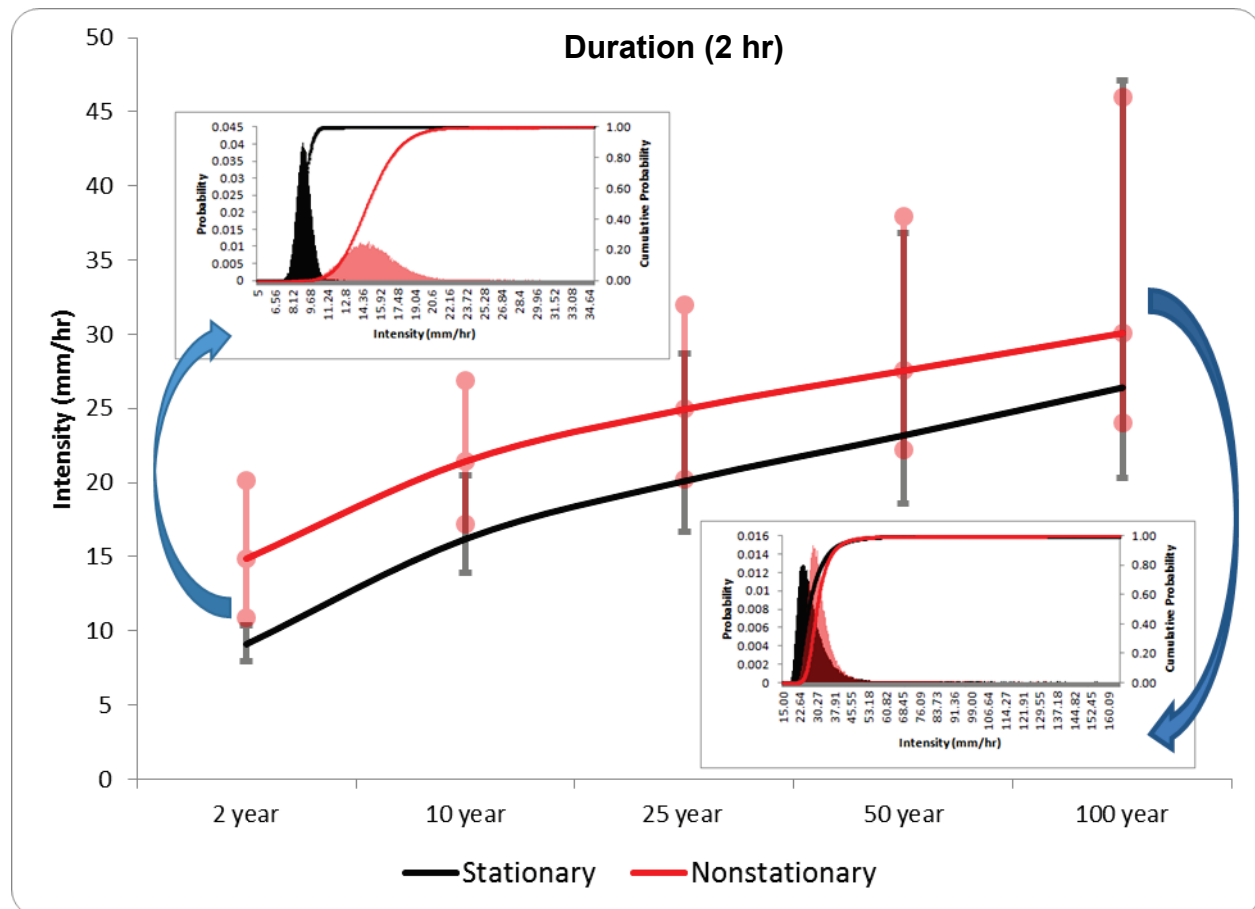


Figure 2. Bayesian MCMC simulation-derived, 2 hr IDF curves for the White Sands National Monument rainfall station computed under stationary and nonstationary conditions. The 2.5, 50, and 97.5 percentile values from each respective predictive posterior distribution are shown at each quantile level. Plots of the posterior predictive distributions for the 2-year and 100-year return period level quantiles are also shown.

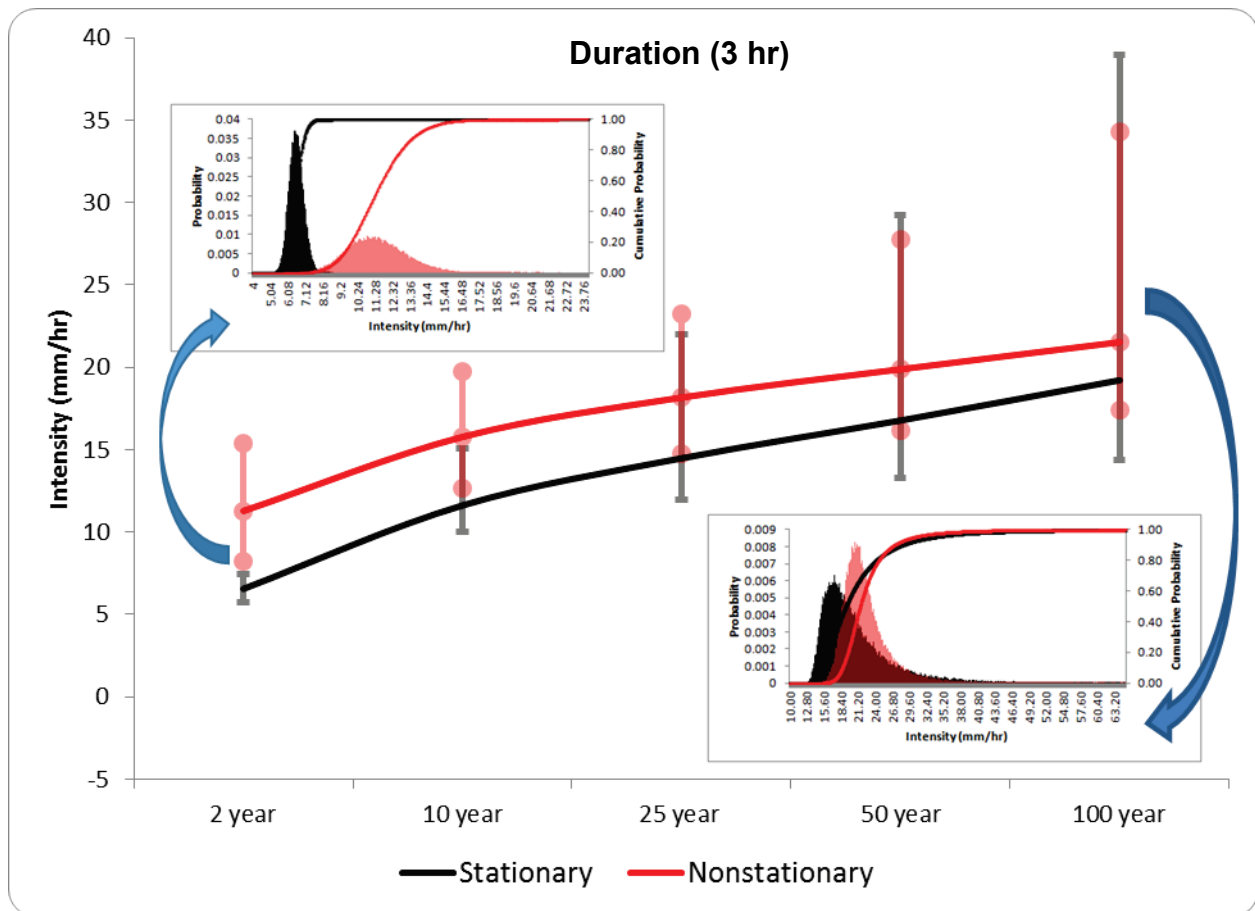


Figure 3. Bayesian MCMC simulation-derived, 3 hr IDF curves for the White Sands National Monument rainfall station computed under stationary and nonstationary conditions. The 2.5, 50, and 97.5 percentile values from each respective predictive posterior distribution are shown at each quantile level. Plots of the posterior predictive distributions for the 2-year and 100-year return period level quantiles are also shown.

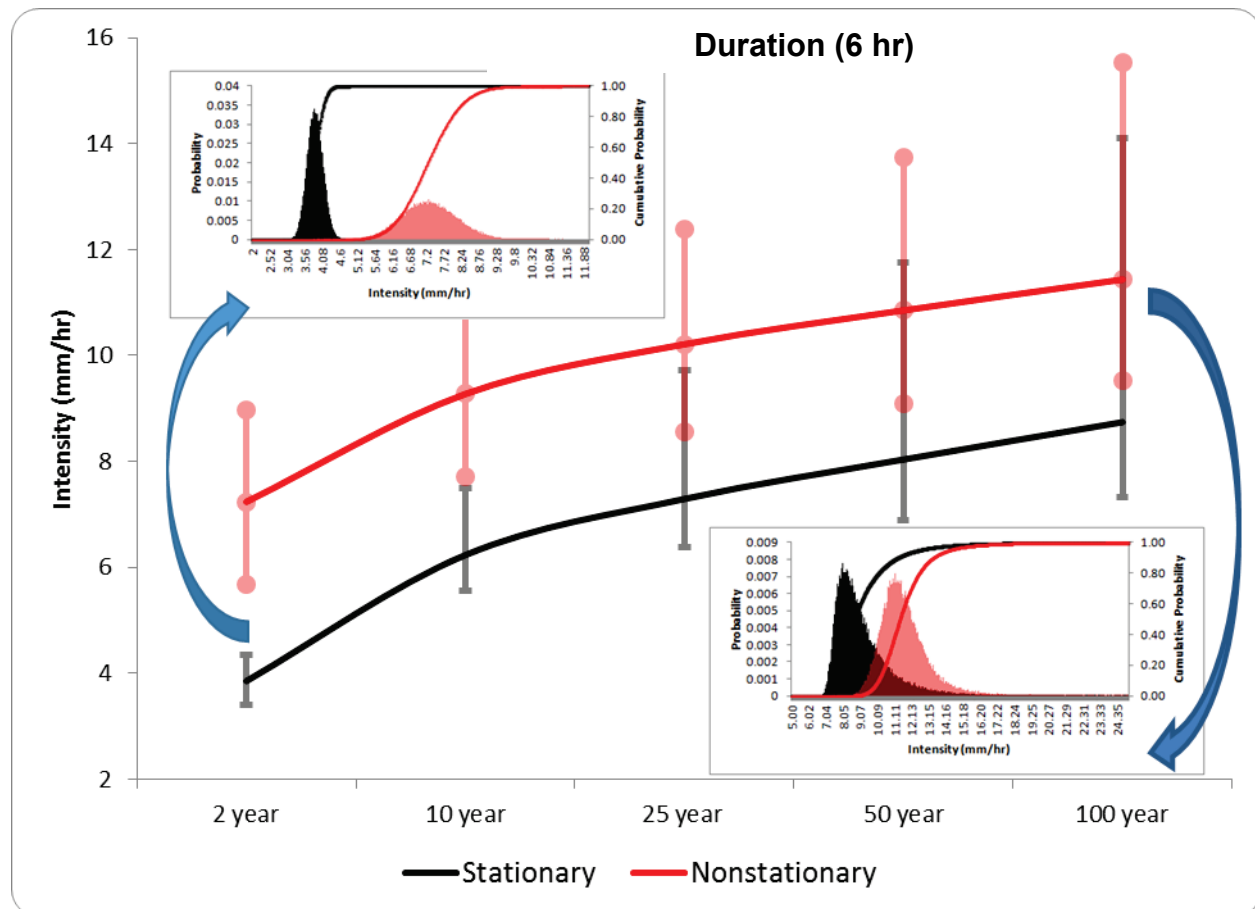


Figure 4. Bayesian MCMC simulation-derived, 6 hr IDF curves for the White Sands National Monument rainfall station computed under stationary and nonstationary conditions. The 2.5, 50, and 97.5 percentile values from each respective predictive posterior distribution are shown at each quantile level. Plots of the posterior predictive distributions for the 2-year and 100-year return period level quantiles are also shown.

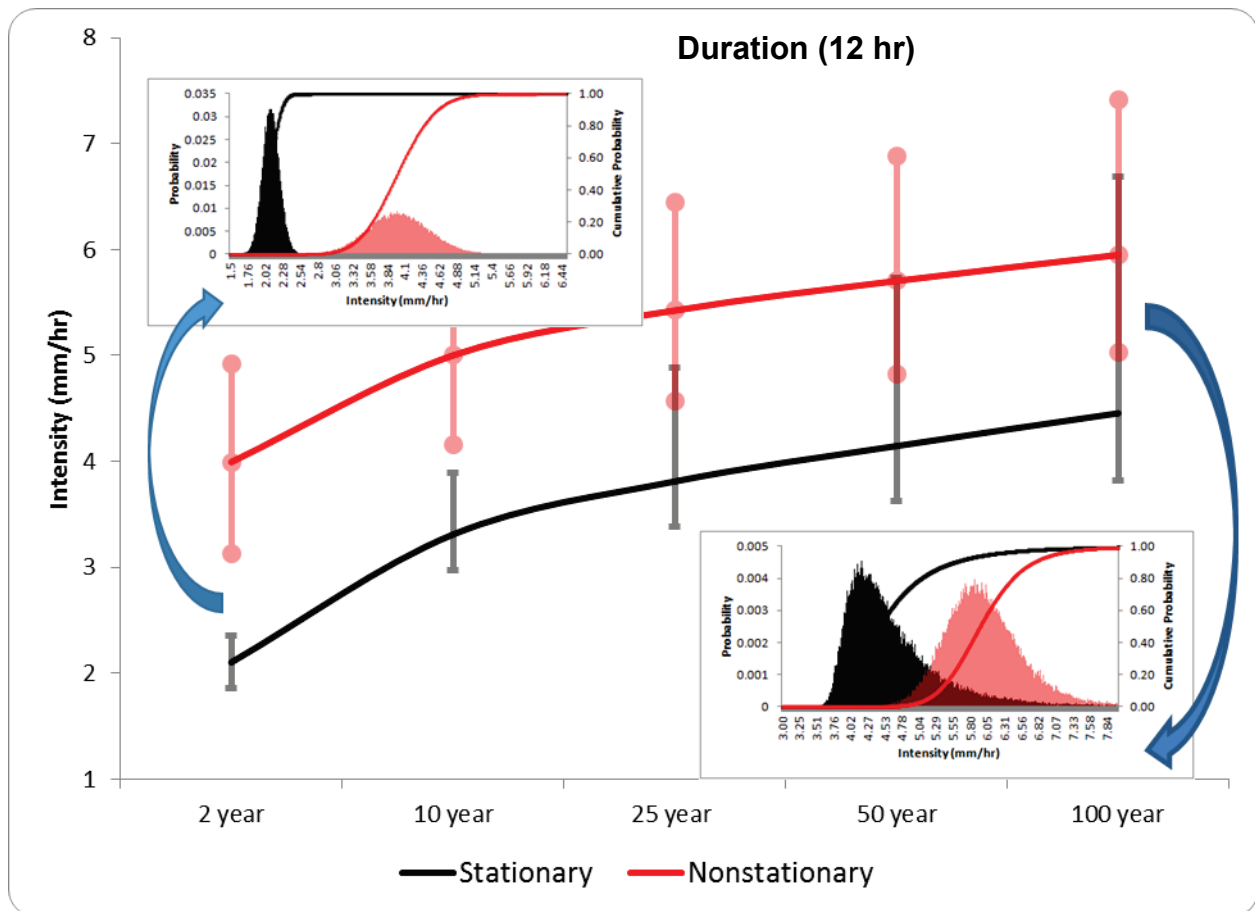


Figure 5. Bayesian MCMC simulation-derived, 12 hr IDF curves for the White Sands National Monument rainfall station computed under stationary and nonstationary conditions. The 2.5, 50, and 97.5 percentile values from each respective predictive posterior distribution are shown at each quantile level. Plots of the posterior predictive distributions for the 2-year and 100-year return period level quantiles are also shown.

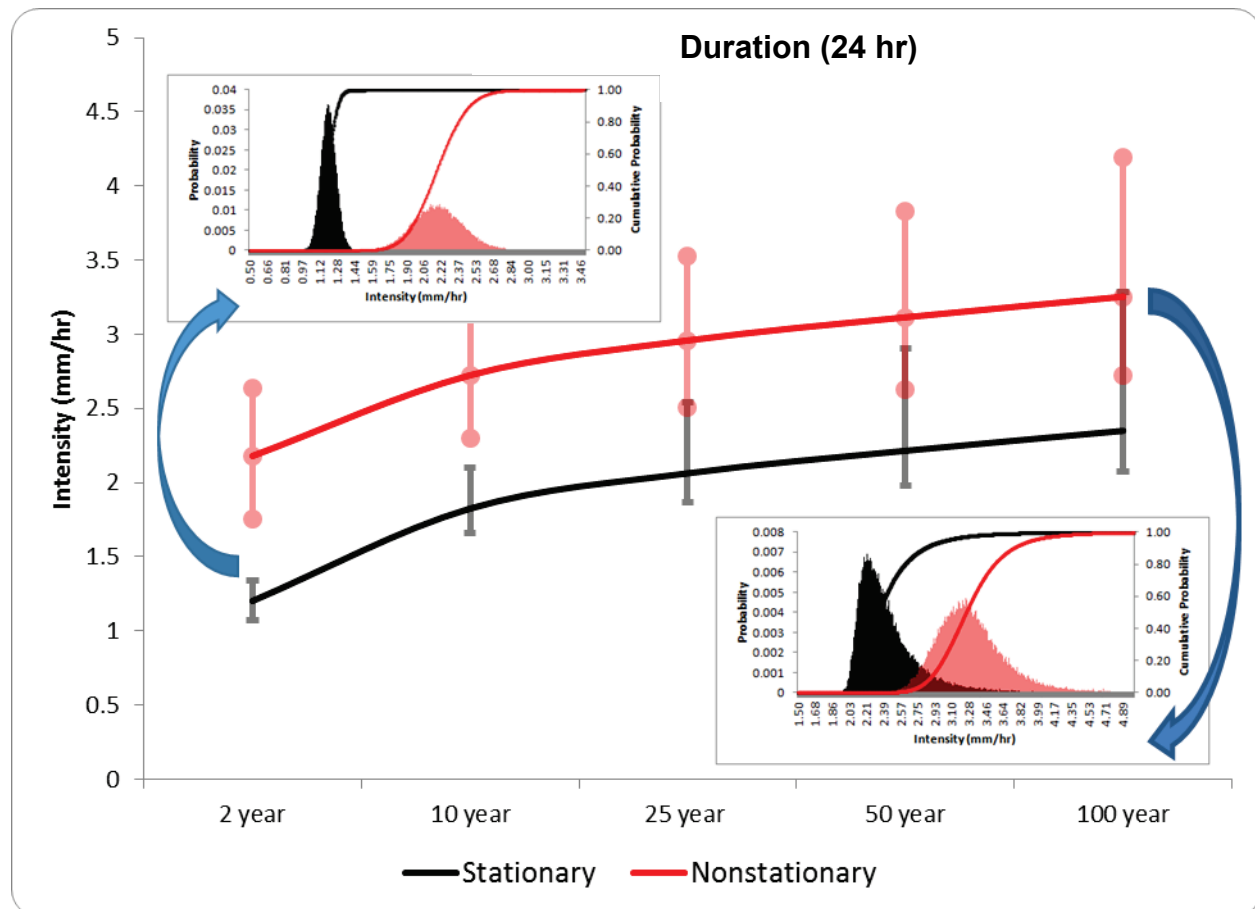


Figure 6. Bayesian MCMC simulation derived, 24 hr IDF curves for the White Sands National Monument rainfall station computed under stationary and nonstationary conditions. The 2.5, 50, and 97.5 percentile values from each respective predictive posterior distribution are shown at each quantile level. Plots of the posterior predictive distributions for the 2-year and 100-year return period level quantiles are also shown.

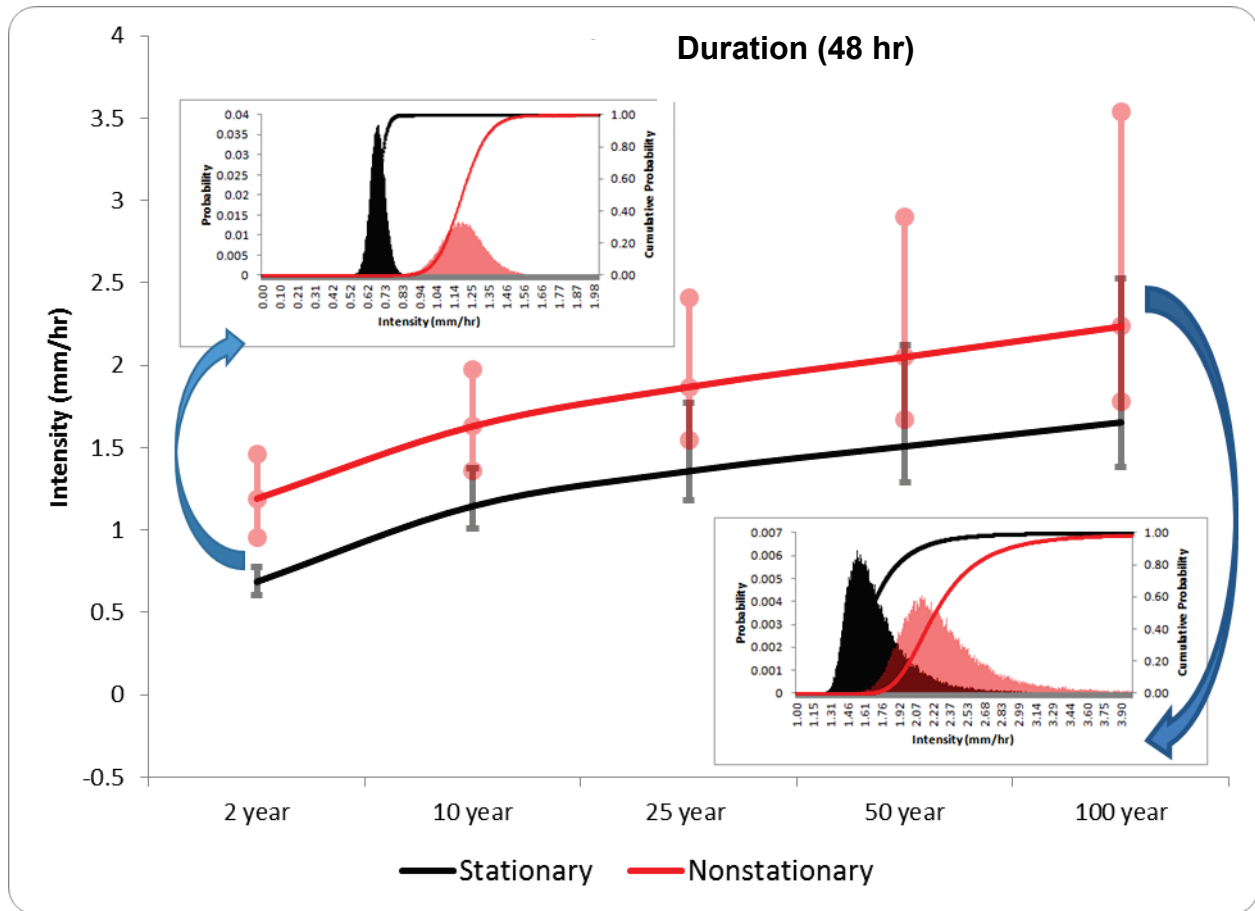


Figure 7. Bayesian MCMC simulation-derived, 48 hr IDF curves for the White Sands National Monument rainfall station computed under stationary and nonstationary conditions. The 2.5, 50, and 97.5 percentile values from each respective predictive posterior distribution are shown at each quantile level. Plots of the posterior predictive distributions for the 2-year and 100-year return period level quantiles are also shown.

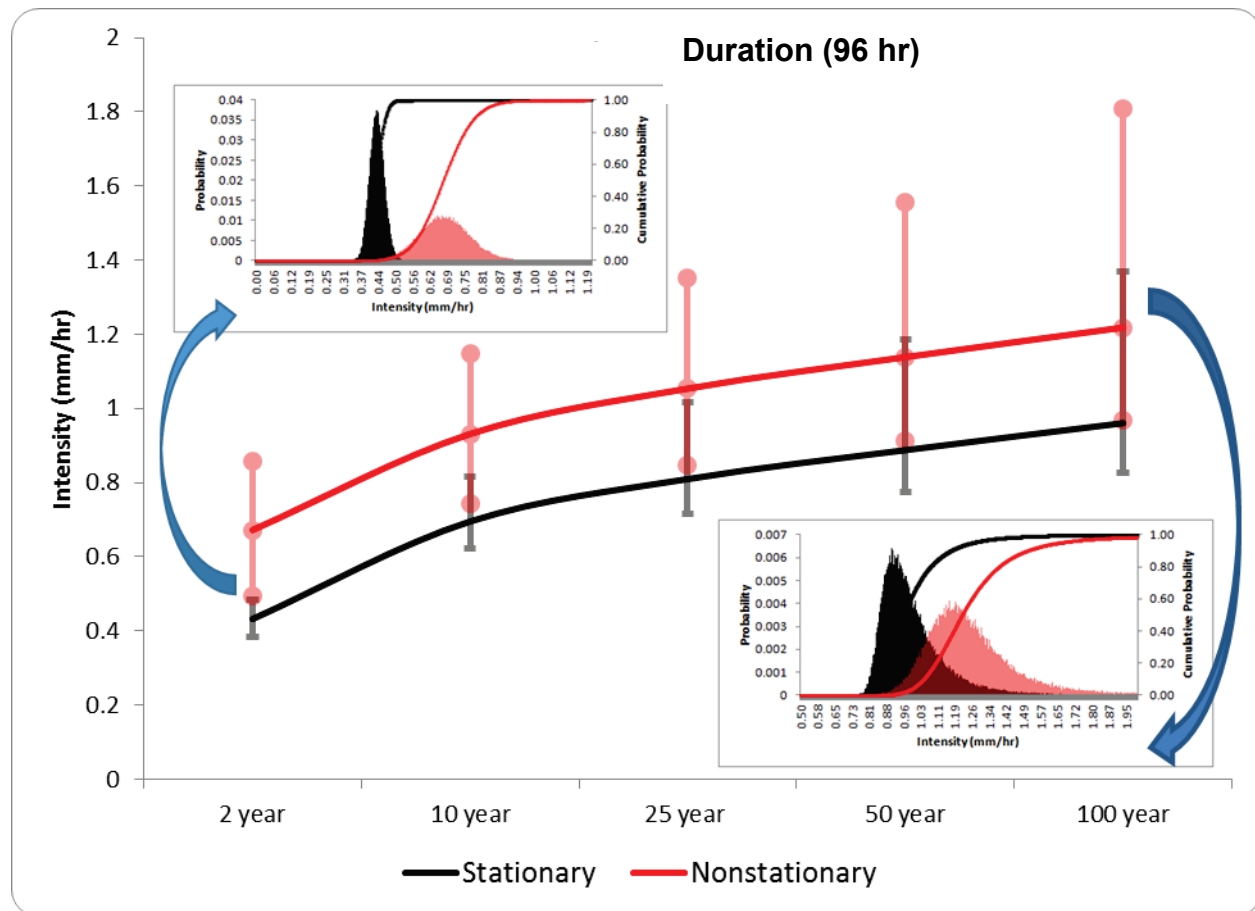


Figure 8. Bayesian MCMC simulation-derived, 96 hr IDF curves for the White Sands National Monument rainfall station computed under stationary and nonstationary conditions. The 2.5, 50, and 97.5 percentile values from each respective predictive posterior distribution are shown at each quantile level. Plots of the posterior predictive distributions for the 2-year and 100-year return period level quantiles are also shown.

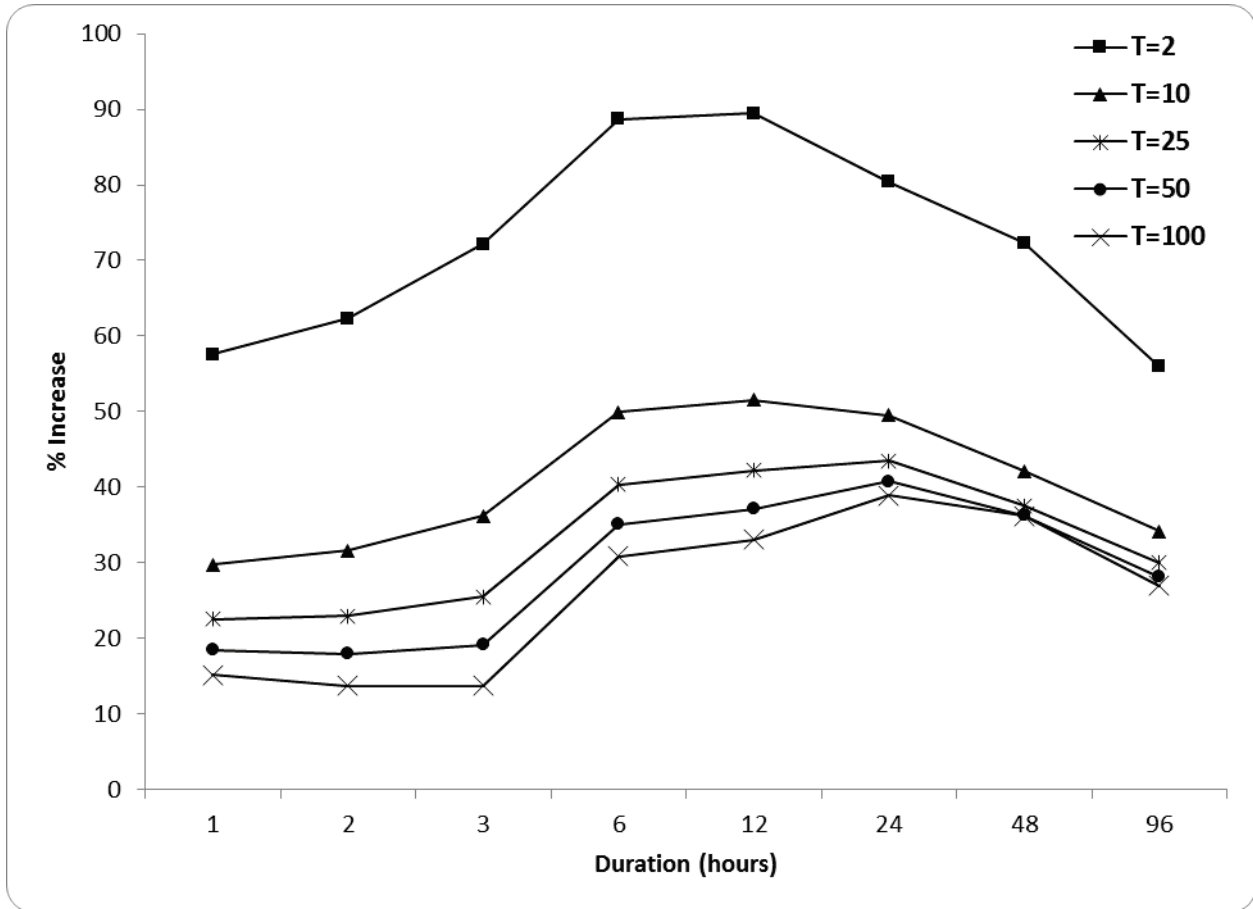


Figure 9. Computed percent increase for five distinct return periods, T , obtained when the nonstationary PM-based quantile estimates are compared with their counterparts that were computed assuming a stationary climate.

DISCUSSION: With the intent to illustrate the potential negative impacts of ignoring nonstationarity at a site that has been assessed otherwise, 16 distinct MCMC simulations were performed to develop IDF curves under stationary and also nonstationary conditions for the White Sands National Monument rainfall station for eight distinct durations. A strength of the MCMC methodology for IDF curve development is the flexibility and ease with which one can incorporate a treatment of a nonstationary climate condition into the analysis. Nonstationarity was treated by specifying the GEV distribution location parameter to vary linearly in time, and as previously mentioned, the results presented in Tables 1–3 and Figures 1–9 are for a projection forward in time. Cheng and AghaKouchak (2014) underscored a primary strength of the MCMC methodology for IDF curve development, viz., its capacity to formally quantify uncertainty for the computed quantiles. This capacity is clearly emphasized graphically in Figures 1–8 wherein the posterior predictive distributions (pdfs and cdfs) for the 2-year and 100-year quantiles, under stationary and also nonstationary conditions, characterized using the post burn-in random draws, are shown for each duration. It is also emphasized in the same set of figures, with the IDF curves themselves including a display of the 95% credible interval together with the 50th percentile value at each quantile level.

Several observations can be made upon examination of the results encapsulated in Tables 1–3 and Figures 1–9 for the White Sands National Monument rainfall station:

1. The stationary assumption delivers IDF curves that underestimate extreme events across all durations and return periods when the comparisons are based on the computed and reported 50th percentile values for each quantile (viz., the red lines are always above the black lines in Figures 1–8).
2. In particular, for example, for a 2-year 2 hr storm, the difference between the nonstationary (14.66 mm/hr) and stationary (9.03 mm/hr) PM extreme precipitation estimates is approximately 5.63 mm/hr (+62.3%) while for a 10-year 1 hr event, the difference between the nonstationary (35.03 mm/hr) and stationary (27.01 mm/hr) PM extreme precipitation estimates is 8.02 mm/hr (+29.7%). These values are both in close agreement with previously reported comparisons (Cheng and AghaKouchak 2014).
3. The most substantial underestimation of extremes that result from ignoring the nonstationary condition occur at the 12 hr duration for the 2-year and 10-year return periods while for the remaining return periods it occurs at the 24 hr duration.
4. The percent increase between the PM nonstationary and stationary extreme precipitation estimates decreases as the return period increases, viz., in Figure 9, the curve for $T=2/T=10/T=25/T=50$ is always above the curve for $T=10/T=25/T=50/T=100$, across all durations.
5. The largest percent increase between the PM nonstationary and stationary extreme precipitation estimates occurs at the 2-year return period level, which is significantly higher than for the remaining return periods.
6. While the differences between the PM nonstationary and stationary estimates decrease as the duration increases, as Cheng and AghaKouchak (2014) observed, the computed percent increases nevertheless indicate notable change occurring across all durations and return periods. In fact, the percent increases are greater for the longer duration events than for the shorter events for all but the 2-year return period.
7. The 95% credible intervals shown at each quantile level suggest that for a given duration, the uncertainty in the computed quantiles for the nonstationary and stationary estimates grow with increasing return period and that this occurrence is more dramatic for the stationary estimates than for the nonstationary estimates. Moreover, for the nonstationary estimates, this phenomenon is less active at the 6 hr, 12 hr, and 24 hr durations wherein the 95% credible intervals are observed to be more uniform across the five return periods relative to the remaining durations.
8. At the 2 hr duration, the 95% credible interval of the stationary 100-year quantile covers its nonstationary counterpart. For the 3 hr duration, the stationary 50-year and 100-year quantile 95% credible intervals cover their nonstationary counterparts.
9. For many durations and return periods, the 50th percentile of stationary simulations are below the lower bounds of the credible intervals of their nonstationary counterparts.
10. Across all durations, the nonstationary 95% credible intervals for the 2-year and 10-year quantile levels are greater than their stationary counterparts.
11. In general, for any given duration, the nonstationary and stationary 95% credible intervals intersect more as the return period increases.
12. For the 2-year return period quantile level, across all durations, the nonstationary and stationary 95% credible intervals intersect minimally, if at all.

13. For a given duration, across all return periods, the degree of intersection of the nonstationary and stationary 95% credible intervals is the least at the 24 hr duration.
14. The posterior predictive distributions presented for the 2-year and 100-year quantiles for each duration, and which are also available for other quantile levels using the available draws from each respective duration-based MCMC simulation, are a means by which to make probabilistic statements regarding extreme precipitation at the White Sands National Monument rainfall station. For example, the approximate cumulative probability that the 3 hr 2-year rainfall intensity, assuming stationarity, is less than or equal to 8 mm/hr is effectively 1 (0.9986); whereas, the approximate complement cumulative probability for the same intensity, duration, and frequency computed under a nonstationary climate condition is 0.983. The approximate cumulative probability that the 24 hr 100-year rainfall intensity, assuming stationarity, is less than or equal to 2.75 mm/hr is 0.873; whereas, the approximate complement cumulative probability for the same intensity, duration, and frequency computed under a nonstationary climate condition is 0.968.

Cheng and AghaKouchak (2014) mentioned that potential nonuniform and climate-induced changes on heavy rainfall events call into question the accuracy and adequacy of current infrastructure design concepts, which rely on an assumption of climate stationarity. Those comments are reinforced herein with a comprehensive revisit of a Bayesian-based comparative evaluation, under stationary and nonstationary conditions, of the NOAA Atlas 14 update extreme rainfall dataset associated with the White Sands National Monument rainfall station. This extreme rainfall dataset was assessed, via the application of statistical trends tests, to exhibit nonstationary behavior. The IDF curves developed under stationary and nonstationary climate conditions using Bayesian MCMC clearly indicate that a stationary assumption underestimates extreme rainfall at the White Sands National Monument rainfall station, across all durations and return periods, and particularly for the 2-year return period and the 6 hr and 12 hr durations. The results were presented in a manner to communicate to the reader that the Bayesian methodology for IDF curve development profiled herein provides one with formal estimates of uncertainty, to support risk-informed hydrologic analysis.

CONCLUSIONS AND FUTURE WORK: The content of this document has demonstrated a new USACE capacity to develop IDF curves under stationary and nonstationary climate conditions using Bayesian MCMC simulation by independently revisiting parts of the example originally profiled by Cheng and AghaKouchak (2014) for the White Sands National Monument rainfall station located in the state of New Mexico. A Bayesian analysis is attractive in that it permits one to flexibly and coherently treat a nonstationary analysis. Moreover, its application provides a basis to make formal probabilistic-based inferences regarding the rainfall quantiles. A Bayesian analysis satisfies the requirements of existing USACE policy guidance regarding flood damage reductions studies, viz., a probabilistic analysis of “all key variables, parameters, and components of flood damage reduction studies.” For example, with a given formal Bayesian-supervised IDF curve development analysis, each post MCMC sampler burn-in random draw would define a unique IDF curve. A design storm hyetograph can be developed from each IDF curve and subsequently used to force a calibrated and validated precipitation-runoff model. Probability-based, risk-informed hydrologic analysis and design, including the possible consideration and treatment of a nonstationary climate condition, could be supported by performing such a hydrologic modeling analysis across all of the generated post burn-in draws. The process would yield an empirical predictive distribution for a hydrologic state (e.g., discharge at a given location).

In this Technical Note, multiple MCMC duration-based simulations were performed using the NOAA Atlas 14 update dataset associated with the White Sands National Monument rainfall station to simultaneously optimize and infer the posterior distribution of the generalized extreme value distribution parameters implied by each Bayesian modeling analysis, performed under stationary and also nonstationary climate conditions. It is underscored to the reader for clarity that distributions other than the GEV could easily be considered within the Bayesian analysis framework. The results obtained, processed, and presented in tabular form and also graphically clearly emphasize that ignoring the nonstationary assumption could lead to substantial underestimation of rainfall extremes.

The Bayesian-based analysis methodology profiled herein is attractive because it dovetails with the stated goal of existing related USACE policy guidance. In particular, the USACE is required to perform risk and uncertainty analyses in the process of planning, design, and operation of all civil works flood risk management projects as described in Engineer Regulation (ER) 1105-2-100 (2006), and its cited references (e.g., Engineer Manual 1110-2-1619 [1996]). The risk-informed analysis framework presented in ER 1105-2-100 (2006) jointly promulgated by the USACE Planning and Engineering communities of practice, requires acknowledgement of and accounting for error and uncertainty in the “key variables, factors, parameters, and data components” relevant to the planning and design of flood damage reduction projects. By capturing and quantifying “the extent of the risk and uncertainty in the various planning and design components of an investment project,” it permits for an evaluation of the tradeoff between risks and costs. The ultimate goal of the policy guidance is probabilistic analysis of “all key variables, parameters, and components of flood damage reduction studies.” Bayesian MCMC simulation for IDF curve development is also attractive because it flexibly permits for the evaluation of a nonstationary climate condition. For the analysis considered herein, the nonstationary climate condition was treated by allowing the GEV distribution location parameter to vary linearly in time. Future related USACE research and development will consider and compare additional temporal treatments of the distribution parameters. For example, of interest, is to further explore and understand the basis for the previously mentioned observed differences, at the White Sands National Monument rainfall station, at and across each quantile level, for the nonstationary uncertainty estimates at a given duration in comparison with their stationary counterparts. Other approaches such as the so-called “effective return levels” can also be explored for treating nonstationarity.

The station-specific Bayesian MCMC approach profiled herein can be applied on a point-by-point basis to generate, by way of interpolation, spatial maps of hydrometeorological extremes; however, such an approach is not explicitly spatial. An additional alternative approach for updating precipitation frequency estimates under stationary and/or nonstationary climate conditions involves using spatially explicit Bayesian modeling analysis. Spatial Bayesian modeling is flexible in that it can accommodate different covariate relationships, it combines information from different durations, and the sources of uncertainty are easily tracked and quantified from the estimated posterior distributions. Spatial Bayesian modeling analysis, in particular, Bayesian Hierarchical Modeling (BHM), is a flexible and coherent statistical framework for quantifying the uncertainty of IDF estimates, which vary with location and duration. Recent applications of BHM include, among others, Soltyk et al. (2014), Lehman et al. (2013), and Cooley et al. (2007). BHM is another direction for future related USACE research and development, including its interface with model-simulated data from a regional climate model for an alternative Bayesian supervised treatment of a nonstationary climate condition. Moreover, BHM is attractive in that it could be the basis for the

generation of spatial information expansion data to be used in a Bayesian analysis of the flood frequency hydrology concept (Skahill et al., in preparation).

ADDITIONAL INFORMATION: This CHETN was prepared as part of the Extreme Hydrologic Events work unit in the Infrastructure R&D Program and was written by Drs. Brian E. Skahill (Brian.E.Skahill@usace.army.mil) and Aaron Byrd (Aaron.R.Byrd@usace.army.mil) of the U.S. Army Engineer Research and Development Center (ERDC) Coastal and Hydraulics Laboratory (CHL); Dr. Amir AghaKouchak (amir.a@uci.edu) of the University of California at Irvine Hydroclimate Research Laboratory; Dr. Linyin Cheng (linyin.cheng@noaa.gov) of the NOAA Earth Systems Research Laboratory Physical Sciences Division; and Dr. Joseph Kanney (Joseph.Kanney@nrc.gov) of the Nuclear Regulatory Commission Office of Research. The Program Manager is Dr. Cary Talbot, and the Technical Director is William Curtis. This CHETN should be cited as follows:

Skahill, B. E., A. AghaKouchak, L. Cheng, A. R. Byrd, and J. Kanney. 2016. *Bayesian inference of nonstationary precipitation intensity-duration-frequency curves for infrastructure design*. ERDC/CHL CHETN-X-2. Vicksburg, MS: U.S. Army Engineer Research and Development Center. <http://chl.erd.usace.army.mil/chetn>

REFERENCES

- Cheng, L., and A. AghaKouchak. 2014. Nonstationary precipitation intensity-duration-frequency curves for infrastructure design in a changing climate. *Sci. Rep.* 4:7093. doi:10.1038/srep07093
- Cooley, D., D. Nychka, and P. Naveau. 2007. Bayesian spatial modeling of extreme precipitation return levels. *Journal of the American Statistical Association* 102:824–840.
- Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 2004. *Bayesian data analysis, second edition, texts in statistical science*. United Kingdom: Chapman & Hall/CRC.
- Gelman, A., and D. B. Rubin. 1992. Inference from iterative simulation using multiple sequences. *Stat. Sci.* 7:457–472.
- Hastings, W. K. 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 57:97–109.
- Intergovernmental Panel on Climate Change (IPCC). 2007. *Climate change 2007: The physical science basis*. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, edited by S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller. Cambridge, United Kingdom and New York: Cambridge University Press.
- IPCC. 2013. *Climate change 2013: The physical science basis*. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P. M. Midgley. Cambridge, United Kingdom, and New York: Cambridge University Press.
- Koutsoyiannis D., D. Kozonis, and A. Manetas. 1998. A mathematical framework for studying rainfall intensity-duration-frequency relationships. *Journal of Hydrology* 206:118–135.
- Kunkel, K. E., T. R. Karl, D. R. Easterling, K. Redmond, J. Young, X. Yin, and P. Hennon. 2013. Probable maximum precipitation and climate change. *Geophys. Res. Lett.* 40:1402–1408. doi:10.1002/grl.50334
- Lehmann, E. A., A. Phatak, S. Soltyk, J. Chia, R. Lau, and M. Palmer. 2013. Bayesian hierarchical modeling of rainfall extremes. In *MODSIM: 20th International Congress on Modelling and Simulation, Adelaide, 1–6 December 2013*.

- Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. 1953. Equation of state calculations by fast computing machines. *J Chem Phys* 21:1087–1092.
- Overeem, A., I. Holleman, and A. Buishand. 2008. Rainfall depth-duration-frequency curves and their uncertainties. *Journal of Hydrology* 348:124–134.
- Perica, S., D. Martin, S. Pavlovic, I. Roy, M. St. Laurent, C. Trypaluk, D. Unruh, M. Yekta, G. Bonnin. 2013. *NOAA Atlas 14, volume 9, precipitation-frequency atlas of the United States, Southeastern States*. Silver Spring, MD: NOAA, National Weather Service.
- Skahill, B. E., A. Viglione, and A. R. Byrd. In preparation. *A Bayesian analysis of the flood frequency hydrology concept*. ERDC/CHL Technical Note. Vicksburg, MS: U.S. Army Engineer Research and Development Center.
- Soltyk, S., M. Leonard, A. Phatak, and E. Lehmann. 2014. Statistical modelling of rainfall intensity-frequency-duration curves using regional frequency analysis and bayesian hierarchical modelling. In *Hydrology and Water Resources Symposium 2014*, 302–309. Barton, ACT: Engineers Australia.
- Srivastav, R., A. Schardong, and S. Simonovic. 2014. Equidistance quantile matching method for updating IDF curves under climate change. *Water Resources Management* 28(9):2539–2562.
- Stedinger, J. R., R. M. Vogel, and E. Foufoula-Georgiou. 1992. Frequency analysis of extreme events. In *Handbook of hydrology*, edited by D. R. Maidment, 18. New York: McGraw-Hill.
- ter Braak, C. J. F. 2006. A Markov Chain Monte Carlo version of the genetic algorithm differential evolution: Easy Bayesian computing for real parameter spaces. *Stat. Comput.* 16:239–249.
- ter Braak, C. J. F., and J. A. Vrugt. 2008. Differential evolution Markov chain with snooker updater and fewer chains. *Stat. Comput.* 18(4):435–446. doi: 10.1007/s11222-008-9104-9
- Trenberth, K. E. 2011. Changes in precipitation with climate change. *Clim Res* 47:123.
- U.S. Army Corps of Engineers (USACE). 2006. Risk based analysis for evaluation of hydrology/hydraulics, geotechnical stability, and economics in flood damage reduction studies. ER 1105-2-101. Washington, DC.
- _____. 1996. Risk-based analysis of flood damage reduction studies. EM 1110-2-1619. Washington, DC.
- Yilmaz, A. G., I. Hossain, and B. J. C. Perera. 2014. Effect of climate change and variability on extreme rainfall intensity-frequency-duration relationships: A case study of Melbourne. *Hydrology and Earth System Sciences* 18:4065–4076.

NOTE: The contents of this technical note are not to be used for advertising, publication, or promotional purposes. Citation of trade names does not constitute an official endorsement or approval of the use of such products.